DOCUMENT RESUME

ED 299 283 TM 011 976

AUTHOR Morra, Sergio; Scopesi, Alda
TITLE On the Validity of Working Memory Measures for Elementary School Children.
PUB DATE [85]
NOTE 47p.; Tables and portions of text contain light, broken print.
PUB TYPE Reports - Research/Technical (143)

EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS *Cognitive Tests; Elementary Education; *Elementary School Students; *Memory; Spatial Ability; Speech Tests; *Test Validity; Verbal Ability; Verbal Tests
IDENTIFIERS *Working Memory

ABSTRACT

The literature on the measurement of working memory is discussed, with emphasis on the validity of such measurement with elementary school students. Approaches by J. Pascual-Leone, R. Case, A. D. Baddeley, and C. Hulme are considered. An experimental study is presented, in which 100 boys and 91 girls (aged from 6 years to 10 years 11 months) were given a battery of 17 tests in four sessions, including well-known tests of verbal and spatial ability and measures that have already been proposed for various aspects of working memory. Eight subjects were excluded from some of the analyses because they made too many errors on the letter classification test. Three tests were group-administered. Two group sessions were followed by two individual sessions. The statistical analyses included factor analysis, partial correlations, and analysis of variance. The results show that five tests can be used in the measurement of the "central components" of working memory; this is consistent with the theories of Pascual-Leone and Case. On the other hand, Hulme's hypothesis that the speed of articulation of school children may predict the capacity of a "peripheral" component (i.e., the articulatory loop) seems to be disproved. Five tables summarize test data, and a 68-item reference list is provided. (SLD)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

Sergio MORRA, Alda SCOPESI

ED299283

ON THE VALIDITY OF WORKING MEMORY MEASURES FOR ELEMENTARY SCHOOL CHILDREN.

Authors' affiliation:
Dr. S.Morra, Università di Padova
Dr. A.Scopesi, Università di Genova

Mailing address:
Dr. S.Morra
Dipartimento di Psicologia
P.Capitaniato 3
35139 PADOVA (Italy)

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

X This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

SERPIO MORRA

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

TM011976

BEST COPY AVAILABLE

ABSTRACT

The literature on the measurement of working memory is discussed, with emphasis on its measurement in children. An experimental study is presented, with 191 subjects (aged from 6 years to 10 years 11 months) who were given a battery of 17 tests, including well-known tests of verbal and spatial ability and measures which have already proposed for various aspects of working memory. Pascual-Leone's, Case's, Baddeley's and Hulme's approaches are considered.

The statistical analyses include factor analysis, partial correlations, analysis of variance. The results show that there is a group of five tests that can be employed in the measurement of the "central components" of working memory; this is consistent with the theories of Pascual-Leone and Case. On the other hand, Hulme's hypotnesis that the speed of articulation of school children may predict the capacity of a "peripheral" component, i.e. the articulatory loop, seems to be disproved.

Is "working memory" a commonplace notion or a mysterious object? On one side, the idea of a memory system which not only retains information but also allows it to be transformed, re-combined, integrated and worked on, has been proposed in various forms by numerous authors (e.g. Atkinson & Shiffrin, 1968; Baddeley & Hitch, 1974; Bjork, 1975; Craik & Lockhart, 1972; Pascual-Leone, 1970; Posner & Rossman, 1965), and in this sense it is a commonplace idea, which has been very popular for over twenty years. However, the interpretations of this concept and the controversies about it form a truly tangled knot. Some authors (e.g. Allport, 1980; Keating, 1984; Neisser, 1976) even consider implausible the very existence of a central processor of limited capacity. Baddeley (1981, 1986) proposes a componential approach and Case (1985) a unitary conception of working memory. There is also discussion about the role of maturation and experience in its development (see Case, 1985; Chi, 1978; Huttenlocher & Burke, 1976; Pascual-Leone & Goodman, 1979). The attempts to disentangle this knot give rise to some of the most stimulating questions in the field of cognitive research.

One crucial point is measurement. Each theory which posits the existence of a working memory of limited capacity (or any similar construct) comes up against the problem of measuring the limits of this capacity. A true theory calls for adequate instruments of measurement, and a false theory

may encounter anomalies also in validating instruments of measurement.

Some authors have over-simplified the problem of measurement, suggesting that the verbal short-term memory span can be an adequate measure of the limits of capacity. Hypotheses of this type were still being advanced recently (e.g. by Brainerd, 1983 and by Halford, Bain & Mayberry, 1984). When this hypothesis is given close examination, however, experimental results rule it out (Halford & al., 1984; Brainerd & Kingsma, 1985). In fact, digit span and similar tasks do not call for the transformation of input, but merely its reproduction (Case, 1972). It is possible that the simple retention of verbal information does not engage the whole system of working memory, but only a subsystem, which Baddeley & Hitch (1974) have called the articulatory loop. It is also posible that digit span depends too much on the ability to retrieve decaying information (Burtis, 1982). On the distinction between working memory and short-term span, see also Baddeley (1976), Klapp, Mashburne & Lester (1983).

A different approach is based on indirect techniques, such as the measurement of "residual capacity". This is based on the assumption that the mental workload imposed by task A can be measured by the decrease in efficiency in the performance of task B (e.g. of a simple reaction time task) carried out at the same time as task A, compared to when task B is carried out on its own. However, measures of

residual capacity proved to be inadequate even with adult subjects. McLeod (1978) provides a clear experimental demonstration of their unreliability (see also Pew, 1979, for a critical discussion). It should, incidentally, be noted that, although Baddeley's research group often employ techniques based on the simultaneous performance of two tasks in order to study working memory, they have not the aim of obtaining measures of capacity. It is therefore necessary to define and validate techniques of measurement that do not depend on a secondary task. Furthermore, as we have seen, such measures cannot consist merely of tests of short-term memory span.

The problem we examine in this article is precisely that of measurement in case of children. We shall consider measures which are applicable to school children and have been conceived within different theoretical frameworks (cf. Baddeley, 1981; Case, 1985; Pascual-Leone, 1970). The applied aspect of our study is a contribution to determine the validity and applicability of those measures. It has, we believe, also a theoretical relevance, since it tests some predictions derived from such theories.

So far, Baddeley and colleagues have proposed four or five components of working memory: a Central Executive, an Articulatory Loop, a Phonological Store, a Spatial Working Memory, and perhaps a visual imagery component, the distinction of which from the spatial component is still in question. On the architecture of working memory according to

Baddeley's theory, see Baddeley (1986), Baddeley & Liebermann (1980), Logie (1986), Salame' & Baddeley (1982), Vallar & Baddeley (1984). It is assumed that each component has a limited capacity, but a quantitative estimate of capacity has been made only for the Articulatory Loop. According to Baddeley, Thomson & Buchanan (1975), the capacity of this mechanism should be equivalent to a sequence of words (or, at any event, of phonemes) that can be pronounced in approximately 1.5 - 2 seconds.

Most of the research carried out by Baddeley and co-workers has studied adult subjects; however, some of their experimental paradigms have been used also with children (Hitch & Halliday, 1983; Hulme, Thomson, Muir & Lawrence, 1984; Hulme & Muir, 1985). In particular, Hulme & al. have obtained a value for the limits of capacity of the articulatory loop, also with subjects aged 8 to 11, which is very close to that obtained by Baddeley & al. (1975). Moreover, the hypothesis has been formulated that the speed of articulation of an individual can provide a value for the efficiency of the articulatory loop: both Baddeley & al. and Hulme & al. obtain correlations of about .7 between speed of articulation and verbal short-term memory. In our research, therefore, we consider measures of verbal short-term memory and speed of articulation.

A quite different approach was taken by Pascual-Leone, Case, and their collaborators, and numerous experimental studies on children were carried out within their neo-

Piagetian frameworks. The construct [illegible]e is called "central computing space" or "M operator" [illegible]scual-Leone, and "short-term storage space" or "[illegible]" by Case. Since these authors agree substant[illegible] on [illegible]mount of capacity of this mechanism in school child[illegible] we shall consider their theories as equivalent for th[illegible]urpose of this study (with the exception of one aspect that will be considered in the final discussion). Using the Piagetian construct of "scheme" as a unit of measure[illegible] and distinguishing representational schemes (both operative and figurative) from the simpler sensori-motor [illegible]emes, Pascual-Leone gives a value of 1 representational scheme for the capacity of the "M operator" at the age of three, 2 around the age of five, and thus increasing by one unit every two years, finally reaching a capacity equivalent to 7 schemes at the age of fifteen for the average subject. It is no accident that this capacity coincides with Miller's (1956) "magical number". For further details and discussions of the capacity of such a system, see Pascual-Leone (1970, 1978); Case (1972, 1974, 1985); Case & Khanna (1981); Case, Kurland & Goldberg (1982); Scardamalia (1977); Burtis (1982); De Ribaupierre & Pascual-Leone (1979).

There may be some problem in comparing the theories discussed above, not only because of their structural and conceptual differences, but especially for epistemological reasons: Baddeley emphasizes a "bottom-up" empiricist methodology in the construction of his theory, while

BEST COPY AVAILABLE

Pascual-Leone mantains that a rationalist epistemology can give more generality to his theory (see Pascual-Leone & Smirkman, 1980). However, some parts of these theories are not incompatible. For instance, the mechanisms called "M operator" and "I operator" (De Ribaupierre, 1983; Pascual-Leone, 1984; Pascual-Leone & Goodman, 1979; Pascual-Leone & al., 1978, 1981) present substantial analogies with the functions of the Supervisory Attentional System or the Central Executive (Shallice, 1982; Norman & Shallice, 1985; Baddeley & Wilson, 1985; Baddeley, 1986). These last authors consider it premature to attempt to measure such very central mechanisms, while the Canadian research groups have been working for several years on the measurement of the M operator. Thus, the validity of the proposed measures can be one ground for comparing these theoretical approaches. On the other side, Baddeley's articulatory loop could be re-interpreted by the neo-Piagetian authors as a strategy, i.e. an operative scheme; but some aspects of it, such as being time-limited and almost independent from central control devices, could hardly be reconciled with the neo-Piagetian point of view. Therefore, the issues of measurement of capacity of the central component of working memory, and of the articulatory loop, seem very important in comparing these theories.

Various studies have already proposed and in part validated instruments for measuring the capacity at school age of resources such as the M operator or equivalent

constructs (Case, 1985; Globerson, 1983; Johnson, 1982; Miller & Pascual-Leone, 1981; Pascual-Leone, 1970, 1978) or the articulatory loop (Hulme & al., 1984; Hulme & Muir, 1985). Generally, the studies cited have obtained promising results, but none of them sets out to check the independence of the measures proposed from other measures of ability which have been validated in traditional psychometric research. Often correlations between various measures of the same construct are presented, or statistical analyses of progress with age in tests of this kind. But rarely is an attempt made to demonstrate that the tests used are independent, or at least distinguishable, from the classic measures of verbal ability, spatial ability etc. which psychometricians have developed over the course of decades. The problem seems particularly delicate with regard to the measures of the M operator. In fact, if there is such a thing as a central computing space (or a central component of working memory), then its limits of capacity might influence performance in various ways in numerous ability tests (possibly even in all of them), and the consequence would be the virtual impossibility of devising valid tests of this construct. Alternatively, valid tests could be constructed if it is possible to distinguish on theoretical and empirical grounds the cognitive tasks that make considerable use of the central mechanisms from those which make less use of them (see Pascual-Leone, 1969, 1978, 1980).

Case and Globerson (1974) and Globerson (1983, 1985) have started to confront this problem and have at least distinguished the measures of central computing space from those of field dependence (see Witkin & al., 1962). Nevertheless, only one of these studies (Globerson, 1983) considers an appreciable number of measures (ten), that are sufficiently heterogeneous. We carried out a factor analysis of the matrix of correlations presented there, and obtained two factors, the second of which loads just 3 tests of M capacity (CSVI, SSVI and Backward Digit Span). The result is promising, yet should only be considered as a first step; even ten variables are not very many. Furthermore, for reasons inherent in the experimental design, Globerson's vast sample is composed of groups which are very distinct in terms of age and socio-economic status; this stratification may create problems with respect to the assumptions upon which the coefficient of correlation is based. At this point, we felt it necessary to undertake a study which considers a wider range of abilities and tests, with a sufficiently broad and diversified sample, within which, however, the distributions of scores would be unimodal and not too asymmetric.

A final source of inspiration for our research is to be found in the methodological suggestions made by Keating & al. (1985). In comparison with other work relating the correlational study of individual differences to experimental research and to the cognitive theoretical

framework (e.g. Carroll, 1976; De Ribaupierre & Pascual-Leone, 1984; Sternberg, 1981), Keating & al. particularly examine the need to define strict criteria for the empirical validation of constructs, and delineate a series of methodological rules, which for reasons of space cannot be adequately discussed here. Although our research is especially concerned with measures of accuracy (rather than temporal parameters), the reader familiar with Keating's research programme will probably recognize the influence of his methodological approach on our work.

In our research we consider four tests taken from among those which have already been proposed for the measurement of the M operator or working memory (in the sense understood by Case): the Backward Digit Span, the Counting Span Test, the Mr.Cucumber Test, the Figural Intersections Test, and two measures which, by analogy, we also considered worthy of being proposed: Backward Word Span and Corsi's Test Backward. We also consider three measures that refer to the articulatory loop (Forward Word Span, Forward Digit Span, Speed of Articulation). In addition, we use well-known spatial tests (Raven's PM, WISC Block Design, WISC Figure Reconstruction, Goodenough's Draw-a-man, Corsi's Test Forward) and well-known verbal tests (WISC Vocabulary, PMA Verbal Fluency). Finally, we use a variant of Posner's letter classification task [1], adapted for children by Keating & Bobbitt (1978).

METHOD

Subjects

The sample comprised 191 children attending elementary school (100 boys, 91 girls; age range 6,0 to 10,11; mean age 8,10). Precisely there were 10 six-year-olds, 39 seven-year-olds, 50 eight-year-olds, 49 nine-year-olds and 43 ten-year-olds. Eight subjects have been excluded from some of the statisics because they made too many errors in one of the test (the letter classification task).

Procedure

Each subject received, in the same order, 17 tests in four sessions. Three of the tests were group-adminisatered. Two group sessions were followed by two individual sessions. The first group session included:

A) Raven's PM47. The score was the number of correct responses.

B) Draw-A-Man (Goodenough's test). The score was given as according to the manual.

The second group session included:

C) Figural Intersections Test (F.I.T.: see Pascual-Leone & Ijaz, in press). A group test of working memory, devised and partially validated by Pascual- Leone and colleagues, which consists in distinguishing the intersection of a set of geometric figures (one of which may also be irrelevant). The level of each item is given by the number of figures

involved in it. The score was expressed by the highest level in which at least 75% of the responses were correct, provided that this criterion had been fulfilled on all the lower levels.

The first individual session included:

D) Posner's task of rapid classification of letters. The subject had to divide up rapidly piles of 20 cards, 10 of which were 'the same' and 10 'different'. After practice there were three timed trials for each condition. No more than one error was allowed in each trial. The NIPI score was given by the difference between the second best time in the Name Identity condition and the second best time in the Physical Identity condition.

E) Forward Digit Span (from the WISC). The score was the highest number of digits correctly repeated in order.

F) Backward Digit Span (from the WISC). The score was the hignest number of digits exactly repeated in reverse order.

G) Corsi's Test Forward (see Cornoldi & Soresi, 1980). The score was given by the number of positions in the longest series which was exactly reproduced.

H) Corsi's Test Backward . The score was given by the number of positions in the longest series that was reproduced exactly in reverse order.

I) Speed of Articulation. The subject was required to count again and again from one to ten, as fast as she or he could, without stopping; the time was taken after five consecutive countings. The better time out of two trials was considered.

J) Verbal Fluency. This test was taken from Thurstone's PMA test; we adapted it for school-age subjects, presenting it orally and with a time limit of 3 minutes.

K) Figure Reconstruction (from the WISC). The raw score attributed as according to the manual's instructions was considered.

The second individual session included:

L) Counting span (see Case, Kurland & Goldberg, 1982; Case, 1985). It consists in the rapid counting of several sets of coloured points and in remembering the number of each set. The level of each item is given by how many sets it comprises. There are three items for each level from 1 to 8; the test is discontinued when the subject failed on all three. The score was given by the highest level on which the subject got at least two of the three items right.

M) Block Design (from the WISC). Since the distribution of the raw scores was highly asymmetric, they were converted into a weighed score (normal distribution in T scores).

N) Forward Word Span. This is a test that we constructed -- on analogy with the forward digit span -- using high frequency words of concrete meaning and of two or three syllables. The score was given by the highest number of words correctly repeated in order.

O) Backward Word Span. As above, but the words were repeated in the reverse order.

P) Mr. Cucumber Test (see De Avila, Havassy & Pascual-Leone, 1974; Case, 1985). The outline of an 'extraterrestrial'

figure, to which coloured stickers had been attached, is exposed for 5 seconds. The subject then has to indicate on an outline devoid of colours the positions of the stickers. There are three items for each level from 1 to 8; the test is discontinued when the subject fails on all three. One point was given for each consecutive level on which the subject got at least two out of three items right, and one third of a point for each item beyond that level; the score was rounded up or down to the nearest unit.

Q) Vocabulary (from the WISC). The raw score was assigned as according to the manual.

The assignement of subjects to testers in each session, and the hour of testing, could not be systematically controlled, but depended on the momentary possibilities and on the needs of the schools. Error variance may be slightly inflated for these reasons, but we feel that the number of subjects is large enough to obtain all the same significant results, notwithstanding this undesirable feature.

Hypotheses and design

The questions we asked, and the statistical techniques we used in order to answer them, were the following:

1. Is it possible to identify a group of measures of M capacity (i.e., capacity of the central component of working memory), that includes at least Backward Digit Span, Figural Intersections Test, Mr.Cucumber Test and Counting Span? Factor analysis can be used to answer this question.

2. If such a factor emerges, on the common variance among the tests with high loadings in this factor be accounted for, independently of whether the material used in them be of spatial or verbal nature? This hypothesis can be verified by checking whether the "verbal" and "spatial" tests of M capacity are more correlated among themselves than the "verbal" tests of M capacity with the other tests of spatial ability, or the "spatial" tests of M capacity with the other tests of verbal ability. The technique for such comparison will be clarified in the Results section.

3. If a group of tests of M capacity can be distinguished, how do their scores increase with age? Do these tests yield equivalent measures, i.e. are they of equal difficulty? If this is not the case, is the increase in score with age substantially the same for the various tests? An Analysis of Variance with age-groups as "between" factor and tests as a "within" factor can answer these questions. If the tests in question were "pure" measures of a central capacity, an Age x Test ANOVA would prove significant only for the Age factor. Vice versa, very high differences between scores on the various tests would show that the measures are not pure, but reflect the variety of content, of processing strategies, of scoring criteria or yet other factors: therefore, they would not be completely equivalent scales for the measurement of the same capacity. Finally, a significant Test x Age interaction, showing that the scores in the tests increased at a different rate with

age, would be critical evidence against using them as measures of the same capacity.

4. Do the tests which are representative of the capacity or efficiency of the Articulatory Loop, i.e. Forward Digit Span, Forward Word Span and Speed of Articulation, tend to form a group and to be distinguishable from all the other tests? If so, ideally, a specific factor should emerge from the factor analysis; however, three tests may not be enough to give rise to a factor. Nevertheless, these three tests should at least show significant correlations, and be more correlated among them than with other tests.

5. Can the Backward Word Span and the Corsi Test Backward be used as tests of M capacity? Thus we speculate, by analogy with the Backward Digit Span. The factor analysis, and eventually the pattern of correlations and the analysis of variance, should answer also this question.

6. What can be the psychometric interest of the NIPI measure derived from Posner's classification task? The factor analysis and the pattern of correlations should answer this question too.

It should be noticed that our study does not address issues of reliability or fidelity of the tests, sometimes addressed by other authors who have studied them. We are only concerned with the problem of validity, i.e. we ask if the tests of M capacity and of the Articulatory Loop actually measure these constructs.

RESULTS

A Factor Analysis (principal components; orthogonal rotation: only the components with eigenvalue greater than 1 are rotated) yields three factors, which explain 43.7% of the total variance (see table 1).

Insert Table 1 about here

In the first factor, which accounts for 20.6% of the variance, the following tests show the highest loadings: Raven's PM47 (.78), Figure Reconstruction (.71), Block Design (.67), Draw-A-Man (.57), F.I.T. (.56), Corsi's Test Backward (.51). It can be interpreted, as a first approximation, as the usual Spatial Ability factor. However, some details suggest that this factor also shows a dimension of field dependence-independence: the four tests with highest loadings were already shown by Witkin & al. (1962) and Case & Globerson (1974) to involve field-restructuring abilities. The Foward version of Corsi's test has a low loading in this factor, while the Backward version has a higher loading: both are spatial in content, but only in the reverse order test are the field effects misleading. And finally, the fact that analytical definitions are evaluated very positively in the Vocabulary test might explain the unforeseen but not negligible (.49) loading of this test in

the first factor. Therefore, we will label this factor Field Independence & Spatial Ability.

The second factor, which accounts for 13.7% of the total variance, clearly identifies Verbal Ability. In fact, it brings together the following tests: Forward Digit Span (.57), Vocabulary (.56), Forward Word Span (.54), Speed of Articulation (.49), Verbal Fluency (.43), Counting Span (.42), Backward Word Span (.39).

In the third factor, the tests with the highest loadings are: Mr.Cucumber (.53), Counting Span (.48), Backward Digit Span (.40), F.I.T. (.38), Backward Word Span (.35). This third factor, which accounts for 9.4% of the total variance, can reasonably be interpreted as M Capacity (i.e., capacity of the M operator or the central component of working memory). Factor analysis is thus able to isolate a factor which loads the four tests listed in Hypothesis 1. The Counting Span has also a substantial loading in the factor of Verbal Ability, and the F.I.T. has a very substantial loading in the Field Independence & Spatial Ability factor. However, three of the four tests have the highest loading precisely in this factor.

Among the tests whose factorial composition we were interested in exploring (see hypotheses 5 and 6), the Backward Word Span may be acceptable, although not excellent, as a test of M capacity, since it has moderate loadings both in the second and third factor. The two parts of the Corsi Test show distinct patterns; at any event,

there is no evidence that the Corsi Test Backward can be used as a test of M capacity. Finally, the NIPI variable does not show any loading greater than .35.

The zero-order correlations between the tests are almost all sufficiently high, but this is obvious, since the subjects of the experiment are children of quite different ages, and all the abilities we are considering develop with age. Some tests show higher linear correlations with age, especially Vocabulary and Draw-A-Man, which indeed assess knowledge and abilities that gradually accumulate. In order to eliminate that portion of common variance simply due to similar patterns of improvement with age, and only focus on the variance due to the knowledge, capacity or ability measured by the tests, we prefer to partial out age. All the following remarks deal with first-order correlations, with age partialled out.

Preliminary analyses (see table 2) deal with the correlations among the tests that have the highest loadings in each factor. The Figural Intersections Test has one of the highest loadings both in the first and in the third factor. It can be seen that :

- the measures of Field Independence and Spatial Ability show the highest partial correlations (median = .37), all significant at the .001 level but two that are significant at the .01 level;

- the measures of Verbal Ability show moderate partial correlations (median = .22), that are all but one significant at the .01 level, and some at the .001;
- the measures of M capacity show moderate partial correlations (median = .25), most of which are significant at the .001 level, and the rest at the .01.

Insert Table 1 about here

These values are not extremely high because of several reasons: age was partialled out, each test has its own specific features, and in addition some variables could not be controlled (as stated in the Procedure section). However, what is important is the pattern of the correlations, and their significance, rather than their absolute values. Other data, not reported in Table 2, show that the correlations among variables with high loadings in different factors are positive but generally lower (median = .20); there are 38 out of 70 that are significant at the .01 level. The Corsi Test Forward has the highest correlations with the Corsi Test Backward ($r = .388$), with the Counting Span ($r = .258$), and with the Figural Intersections Test ($r = .241$), all significant at $p<.001$. The NIPI variable has only two correlations significant at $p<.01$, with the Forward Digit Span ($r = .218$) and with the Counting Span ($r = .213$).

Insert Table 3 about here

We were especially interested in verifying whether, after having partialled out age, it is possible to identify a portion of variance specific to the tests of M Capacity, regardless of whether these use verbal or spatial materials (see Hypothesis 2). In table 3 we report the data relevant for verifying whether this variance can be distinguished from that of Verbal Ability. In order to test this hypothesis, we shall compare each correlation in the left panel of Table 3 with each correlation in the same line of the right panel. For instance, the Mr.Cucumber Test shows a higher correlation with the Backward Digit Span (.231) than with the Forward Digit Span (.092). There are 30 such comparisons; 27 of them point in the hypothesized direction (whichever the amount of the difference) and only 3 in the opposite direction. These figures are highly significant ($p<.001$) with respect to a binomial of parameters 30 and .5; it can thus be concluded that the distinction between Verbal Ability and M Capacity is well grounded.

Insert Table 4 about here

Table 4 reports the relevant data for the same type of distintion between Field Independence & Spatial Ability, and

M Capacity. Here, too, 30 comparisons are possible. Disregarding also in this case the amount of the difference, only 13 comparisons point in the hypothesized direction and 17 point in the opposite direction. Obviously these figures are not significant; it would seem that our Hypothesis 2 is true as regards Verbal Ability, but false as regards Field Independence & Spatial Ability. This apparently anomalous result will be reconsidered in the final discussion. When we pool together all the relevant comparisons, however, we have 40 of them pointing in the hypothesized direction and 20 in the opposite direction; this result is significant ($p<.01$) with respect to a binomial of parameters 60 and .5.

Insert Table 5 about here

The analysis of variance of the tests of M capacity has the aim of checking whether the difficulty of these tests is substantially similar and whether the increase in scores with age is the same for the various tests (see Hypothesis 3). The mean scores are reported in Table 5. The results of the ANOVA indicate that the means of the various age-groups are significantly different, $F(4,186) = 29.06$, $p<.001$, as are the differences between the mean scores in the single tests, $F(4,744) = 16.71$, $p<.001$. The interaction between the two factors, however, is not significant: $F(16,744) = 1.34$, $p>.16$, that is, the increase with age does not follow

different patterns from one test to another, but can be considered homogeneous for all five tests.

Post hoc comparisons between successive age-groups show significant increases for each year of age: from six to seven, $t(47) = 4.781$, $p<.001$; from seven to eight, $t(87) = 2.956$, $p<.004$; from eight to nine, $t(97) = 2.512$, $p<.02$; from nine to ten, $t(90) = 3.741$, $p<.001$. Paired-t comparisons among the various tests of M Capacity show that the Backward Digit Span is significantly easier than each of the other four tests ($p<.001$ in all cases), and moreover that the Backward Word Span is easier than the Counting Span ($p<.001$) and Mr. Cucumber ($p<.004$), and that the Figural Intersections Test is easier than the Counting Span ($p<.003$). Besides the specific content of these tests and the various abilities and types of knowledge involved, it is possible that the varying scoring criteria contribute towards the differences shown.

It could be concluded that:

a) The significant differences between the scores in the various tests indicate that none of them can be regarded as "pure" measures of the central capacity, but,

b) The extent of these differences and the lack of a significant interaction with age indicate that the average score in a battery of tests can provide a reasonable approximation, i.e. one that is not excessively influenced by other abilities that develop at different paces.

More evidence supporting this interpretation comes from a descriptive data, i.e. the intra-individual variance in the scores for various tests. Let us consider an imaginary subject who in the five tests of M Capacity obtains scores of 2, 3, 4, 5, 6. Of course, the mean is 4, but the difference in performance among the tests is such that one could not mantain with confidence that our hypothetical subject has an M Capacity equivalent to 4 schemes. The variance of his performances would in fact be 2.5. Let us consider now another hypothetical subject, whose scores are 3, 3, 4, 4, 5. The variance of his performances would be 0.56, and thus the mean score in this case would be far more reliable. Since among our 191 subjects 71% show a variance less than or equal to 0.56, this is a further indication of the plausibility of the use of these tests as a battery with a single unit of measurement, namely schemes. However, 10% of the subjects have a variance greater than 1, and this may indicate that for a minority of subjects not even the battery as a whole would provide a completely reliable measure.

As far as the Articulatory Loop is concerned, the mean scores of our subjects in Forward Digit Span and Speed of Articulation yield an estimate of its capacity just slightly lower to that of Hulme and colleagues. The value we obtain, 1.24 seconds, is virtually constant across age- groups. However, there is little evidence that this estimate is meaningful. We can observe (see Table 1) that the measures

proposed for the Articulatory Loop, instead of forming a group in themselves, seem to belong to a set of measures of Verbal Ability. In Table 2 we can see that the correlation between Forward Word Span and Speed of Articulation is not even significant once age has been partialled out. Instead, the variable that correlates most with Speed of Articulation is Vocabulary. [2]

The correlation between Forward Digit Span and Forward Word Span remains significant even if, in addition to age, other indexes of Verbal Ability, i.e. Vocabulary and Verbal Fluency, are partialled out: $r(186) = .322$, $p<.001$. But the correlation between Forward Digit Span and Speed of Articulation is not significant once age, Vocabulary and Verbal Fluency have been partialled out: $r(186) = .132$, i.e. less than 2% common variance. Finally, with these partialisations, the correlation between Forward Word Span and Speed of Articulation falls to $r(186) = .003$.

Bob Logie (personal communication, November 1986) has suggested that these results, which seem difficult to reconcile with Hulme's account or with a developmental interpretation of Baddeley's theory (see Baddeley, 1986), could be due to the wide range of the subjects' ages, where there is a mixture of some subjects who use and others who don't use rehearsal strategies. In point of fact, it is known that spontaneous rehearsal is very rare before the ages of seven or eight. In order to verify this possible interpretation, we carried out further analyses which took

into account only the oldest subjects. When subjects older than eight are considered (N=138), the factor analysis yields four factors: Field Independence & Spatial Ability, and M Capacity (just as in the whole sample), plus two verbal factors, one of which loads Vocabulary (.54), Speed of Articulation (-.53), Verbal Fluency (.46), Forward Digit Span (.37), while the other loads only Forward Word Span (.63) and Forward Digit Span (.36). In this last factor, the loading of Speed of Articulation is .11, instead of a negative loading as it should be (since short times should be associate to good span performances). When subjects older than nine are considered (N=92), Speed of Articulation shows zero-order correlations of -.34 with Vocabulary, -.31 with Forward Digit Span, -.25 with Verbal Fluency and only -.19 with Forward Word Span. Therefore, the possible source of artifact pointed out by Logie appears to be ruled out.

DISCUSSION

The general aim of this study was to ascertain whether the capacity of working memory can be measured. This general aim was specified by six hypotheses.

The first hypothesis concerned the possibility of identifying, by means of a factor analysis, a factor which loads a group of measures of M Capacity. The answer is affirmative: the third factor can be interpreted in this sense. However, it should be noted that these tests tend not to be "pure" measures, inasmuch the Counting Span and the Backward Word Span have also a considerable loading in the verbal factor, and the Figural Intersections Test is also highly loaded by the spatial/analitical factor.

A detailed examination of the partial correlations provides converging evidence for the results of the factor analysis, and in particular allows to verify the second hypothesis. Here there is a problematic result. While the distinction (in terms of the pattern of correlations) between M Capacity and Verbal Ability is very neat, the same distinction between M Capacity and Field Independence & Spatial Ability is not clear at all. In order to account for this result, Pascual-Leone (personal communication, August 1986) suggested to use his construct of "meta-subjective formula" (see Pascual-Leone, 1969, in press; Pascual-Leone & Goodman, 1979). Without going into technical detail, we shall discuss only the essential ideas.

Each group of tests identified by factor analysis is characterized by similarities both in the content and in the use of more general psychological mechanisms, such as the M Operator and the other "constructive operators". The first factor would, therefore, refer to tests which have in common both spatial content and a "meta-subjective formula" similar to that of the tests of field-independence, which expresses the need of fully using the M Operator in order to keep activated all the information necessary to overcome misleading field effects. By contrast, the second factor loads tests which have in common not only verbal content but also a "meta-subjective formula" characterized by the possibility of obtaining good performances even when using only in part the capacity of the M Operator, thanks to the use of strategies which are largely learned (rehearsal, use of semantic or other types of links to facilitate retrieval). The third factor loads tests which have in common not only the necessity to use fully the M Operator, but also a processing aspect which could be called "temporal sequences". This is evident in three tests: Counting Span, Backward Digit Span and Backward Word Span; but it also reveals itself in the successive acts of encoding executed in order to analyse appropriately the stimuli in the Mr. Cucumber Test, and perhaps also to identify the single geometric forms in the composite stimuli of the F.I.T. Nevertheless, it is important to note that it is not a question of a factor characterized solely by its temporal

content; indeed, Forward Digit Span, Forward Word Span and the two parts of Corsi's test share the same temporal content, but have lower loadings in this factor.

This interpretation explains why the first and third factors load tests which are so highly correlated with one another. The reason would be that they all make great use of the M Operator. However, this resource would be much taxed in one case, for overcoming an embedding perceptual context, and in the other case, for storing the output of a sequence of previous operations. Since we designed our study treating Pascual-Leone's and Case's theories as equivalent, we did not consider in our hypotheses the possibility that the factor related to measures of M Capacity could be well distinguishable from the verbal factor, but not from the spatial/analytical factor. As this turns out to be the case, the construct of "meta-subjective formula" seems necessary in order to account for this result. From this point of view, the two theories cannot be considered equivalent, as Case has apparently dropped such construct from his theory (e.g. see Case & al., 1982; Case, 1985). Our study seems to suggest that this construct is indeed needed.

Even though they require a considerable use of the M Operator, the tests loaded in the first factor are in general not utilizable in its measurement. This is because their "meta-subjective formula" is rather that of field independence (see Pascual-Leone, in press; Pascual-Leone & Goodman, 1979), and because they are all spatial in content,

as well as for the reason thet they are not constructed in such a way as to be able to identify clearly the distinct pieces of information activated by the M Operator at any moment (except, of course, for the F.I.T., and to a certain extent Raven's PM 47, whose relationship with the M Operator has been analyzed by Bereiter & Scardamalia, 1977). The tests loaded by the third factor, on the other hand, lend themselves better to the measurement of the M Operator. According to Case (1985), tests which require the storage in working memory of the products of a sequence of previous operations, are particularly suited for this purpose.

Further evidence for the use of the measures loaded by the third factor as a battery of tests of M Capacity is given by an analysis of variance of their scores. As specified by our third hypothesis, the lack of a significant Age x Test interaction supports this interpretation. Additional support comes from a descriptive data, i.e. the intra-individual variance. Thus it can be concluded that, both in research and for applied purposes, M Capacity can be validly measured with a battery of tests including Counting Span, Mr.Cucumber, Figural Intersections, Backward Digit Span and Backward Word Span.

Our fourth hypothesis dealt with the measurability of the Articulatory Loop. The partial correlations arouse considerable perplexity. In fact, the correlation between Forward Word Span and Speed of Articulation was not significant, while Speed of Articulation had the highest

correlation with Vocabulary. Given that the measure we adopted for speed of articulation is different from that used by other researchers (e.g. Hulme & al., 1984), we have recalculated the correlations considering, like them, the number of words per second rather than the time taken; however, the overall pattern of results did not change. Also further analyses, which discarded younger subjects who might not spontaneously rehearse, confirm this pattern and, if anything, make it more evident.

These data seem to disprove Hulme's hypothesis about the possibility of measuring the capacity of the Articulatory Loop in school children. It rather seems that the speed of articulation is linked to a more general verbal ability, which also implies articulatory mechanisms and ability in rehearsal, rather than constituting a specific predictor of verbal short-term memory span.

The two final hypotheses regarded single variables. As to the Backward Word Span, which we thought of as a possible measure of M Capacity, one should note that this test has highest loading in the verbal factor, its highest correlation with the Counting Span (also loaded by the second and third factors) and a high correlation also with Vocabulary. This last fact points to a semantic component, which does not influence, for instance, the Backward Digit Span. Nevertheless, although this measure does not seem to be the most efficacious, altogether it seems to be utilizable as one of the tests of M Capacity.

Corsi's Test Backward, on the other hand, proves to be completely unsuited for this purpose, as it is loaded only by the first factor and shows high correlations essentially with the spatial/analytical tests. The inapplicability of Corsi's Test as a test of working memory may derive from the various ways in which the items can be coded. Besides the possibility of coding individual cubes and reconstructing their order backwards (in which case the test could presumably be used as a test of M Capacity), there are other strategies which the subjects can adopt. For example, he or she can code not the individual elements but the gestures which compose the sequence (whereby one element rewer has to be memorized), or again he or she can try to encode a complex, articulate form (and it is likely that the subjects who are more field-independent are better able to follow this strategy).

The results also suggest that the variable NIPI is of little psychometric interest, thus confirming the outcome of recent research. Originally, Keating & Bobbitt (1978) found a correlation between NIPI and Raven's matrixes (while in our data this correlation is very low). Subsequently, however, Keating & al. (1985) have ruled out the idea that several temporal parameters (including the variable NIPI) taken from experiments on information processing could constitute reliable predictors of performance in verbal or spatial tests. Our results also rule out the hypothesis that the variable NIPI can be a measure of automatisation of

processes specifically tied to working memory. In our data, in fact, this variable is loaded by none of the three factors and shows significant correlations with only two variables.

FOOTNOTES

(1) Posner's task is divided in two parts: in the first the pairs of letters are classified as 'same' (e.g. AA) or 'different' (e.g. Bb) based on their physical apparence, in the second based on the name (e.g. Aa same, AB different). According to the assumptions of the subtractive method for the analysis of reaction times (see Bisiacchi, 1979, for a critique), the Name Identity - Physical Identity (NIPI) difference expresses the time taken to recall the names of the letters. The attempts to interpret this variable in psychometric terms (Keating & Bobbitt, 1978; Keating & al., 1985; Morra, 1984) have yielded altogether negative results. But it has been recently suggested (Keating, personal communication, May 1985) that it could be a measure of general automatisation of processes (in the sense given by Case, 1985). We therefore thought it useful to insert it in this study for explorative purposes.

(2) If the speed of articulation is expressed in words per second rather than in seconds, the value of the correlations changes by one or two hundredths, but the overall aspect of the results remains unchanged.

REFERENCES

ALLPORT A.D. (1980) Attention and performance. In: Claxton G. (ed.), Cognitive psychology: new directions. London : Routledge & Kegan Paul. 112-153.

ATKINSON R.C., SHIFFRIN R.M. (1968) Human memory: a proposed system and its control processes. In: Spence K.W. e Spence J.T. (eds.), The psychology of learning and motivation, vol.II, N.York, Academic Press. 89-195.

BADDELEY A.D. (1976) The psychology of memory. N.York : Basic Books.

BADDELEY A.D. (1981) The concept of working memory: a view of its current state and probable future development. Cognition, 10, 17-23.

BADDELEY A.D. (1986) Working Memory, Oxford : Clarendon Press.

BADDELEY A.D., HITCH G.J. (1974) Working memory. In: Bower G. (ed.), The psychology of learning and motivation, vol.VIII, N.York, Academic Press. 47-89.

BADDELEY A.D., LIEBERMAN K. (1980) Spatial working memory. In: Nickerson R.S. (ed.) Attention and performance, vol. VIII, Hillsdale : Erlbaum. 521-539.

BADDELEY A.D., THOMSON N., BUCHANAN M. (1975) Word-length and the structure of short-term memory. J. Verb. Learn. Verb. Beh., 14, 575-589.

BADDELEY A.D., WILSON B. (1985) Frontal amnesia and the dysexecutive syndrome, presented at the Neuropsychology Meeting, Venezia 1985.

BEREITER C., SCARDAMALIA M. (1979) Pascual-Leone's M construct as a link between cognitive-developmental and psychometric concepts of intelligence. Intelligence, 3, 41-63.

BISIACCHI P. (1979) Lo studio dei tempi di reazione in psicologia cognitivista. In: Legrenzi P. (ed.) Realtà e rappresentazione. Firenze : Giunti-Barbera.

BJORK R.A. (1975) Short-term storage: the ordered output of a central processor. In: Restle F., Shiffrin R.M., Castellan N.J., Lindman H.R., Pisoni D.B. (eds.): Cognitive Theory, vol.I. Hillsdale : Erlbaum.

BRAINERD C.J. (1983) Working-memory systems and cognitive development. In: Brainerd C.J. (ed.), Recent advances in cognitive-developmental theory, N.York : Springer-Verlag. 167-236.

BRAINERD C.J., KINGMA J. (1985) On the independence of short-term memory and working memory in cognitive development. Cogn. Psych., 17, 210-247.

BURTIS P.J. (1982) Capacity increase and chunking in the development of short-term memory. J. Exp. Ch. Psych., 34, 387-413.

CARROLL J.B. (1976) Psychometric tests as cognitive tasks: A new "structure of intellect", in Resnick L.B. (ed.) The nature of intelligence, Hillsdale : Erlbaum. 27-56.

CASE R. (1972) Validation of a neo-Piagetian capacity construct. J. Exp. Ch. Psych., 14, 287-302.

CASE R. (1974) Structures and strictures: some functional limitations on the course of cognitive growth. Cogn. Psych., 6, 544-573.

CASE R. (1985) Intellectual development: birth to childhood. N.York : Academic Press.

CASE R., GLOBERSON T. (1974) Field independence and central computing space. Ch. Dev., 45, 772-778.

CASE R., KHANNA F. (1981) The missing links: stages in children's progression from sensori-motor to logical thought. New Direct. for Ch. Devel., 12, 21-32.

CASE R., KURLAND D.M., GOLDBERG J. (1982) Operational efficiency and the growth of short-term memory span. J. Exp. Ch. Psych., 33, 386-404.

CHI M.T.H. (1978) Knowledge structures and memory development. In: Siegler R.S. (ed.), Children's thinking : what develops ? Hillsdale : Erlbaum. 73-96.

CORNOLDI C., SORESI S. (1980). La diagnosi psicologica nelle difficoltà di apprendimento. Pordenone: Erip.

CRAIK F.I.M., LOCKHART R.S. (1972) Levels of processing : a framework for memory research. J. Verb. Learn. Verb. Beh., 11, 671-684.

DE AVILA E., HAVASSY B., PASCUAL-LEONE J. (1976) Mexican-American school children: a neo-Piagetian analysis. Washington : George Washington Univ. Press.

DE RIBAUPIERRE A. (1983) Un modele neo-Piagetian du developpment: la theorie des operateurs constructifs de Pascual-Leone. Cahiers de Psych. Cogn., 3, 327-356.

DE RIBAUPIERRE A., PASCUAL-LEONE J. (1979) Formal operations and M-power: a neo-Piagetian investigation. New Directions in Ch. Devel., 5, 1-43.

DE RIBAUPIERRE A., PASCUAL-LEONE J. (1984) Pour une integration des methodes en psychologie: approches experimentale, psycho-genetique et differentielle. L'Anneeu Psychologique, 84, 227-250.

GLOBERSON T. (1983) Mental capacity and cognitive functioning: developmental and social class differences. Devel. Psych., 19, 225-230.

GLOBERSON T. (1985) Field Dependence/Independence and mental capacity: a developmental approach. Devel. Rev., 5, 261-273.

HALFORD G.S., BAIN J.D., MAYBERRY M.T. (1984) Working memory and representational processes: implications for cognitive development. In: Bouma H., Bouwhuis D.G. (eds.), Attention and performance, vol.X. Hillsdale : Erlbaum. 459-470.

HITCH G.J., HALLIDAY M.S. (1983) Working memory in children. Ph. Trans. R. Soc. Lond., B-302, 325-340.

HULME C., MUIR C. (1985) Developmental changes in speech rate and memory span: a causal relationship? Brit. J. of Devel. Psych., 3, 175-181.

HULME C., THOMSON N., MUIR C., LAWRENCE A. (1984) Speech rate and the development of short-term memory span. J. Exp. Ch. Psych, 38, 241-253.

HUTTENLOTCHER J., BURKE D. (1976) Why does memory span increase with age? Cogn. Psych., 8, 1-31.

KEATING D.P. (1984) The emperor's new clothes: the "new look" in intelligence research. In: Sternberg R.J. (ed.), Advances in the psychology of human intelligence, vol.II,. Hillsdale : Erlbaum. 1-45.

KEATING D.P., BOBBITT B.L. (1978) Individual and developmental differences in cognitive-processing components of mental ability. Ch. Dev., 49, 155-167.

KEATING D.P., LIST J.A., MERRIMAN W.E. (1985) Cognitive processing and cognitive ability: a multivariate validity investigation. Intelligence, 9, 149-170.

KLAPP S.T., MARSHBURN E.A., LESTER P.T. (1983) Short-term memory does not involve the "working memory" of information processing: the demise of a common assumption. J. Exp. Psych.: Gen., 112, 240-264.

JOHNSON J.M. (1982) The Figural Intersections Test (FIT), a measure of mental attentional energy. Technical report, York University., Ontario, Canada.

LOGIE R.H. (1986) Visuo-spatial processing in working memory. Quart. J. Exp. Psych., 38A, 229-247.

MC LEOD P. (1978) Does probe RT measure central processing demand? Quart. J. Exp. Psych., 30, 83-89.

MILLER G.A. (1956) The magical number seven, plus or minus two: some limits on our capacity for processing information. Psych. Rev., 63, 81-97.

MILLER M.S., PASCUAL-LEONE J. (1981) Disconfirming Jensen experimentally: intellectual versus executive-structural deficiency in underperforming low-SES children. Presented at the Biennal Convention of the Soc. for Research in Child Development, Boston.

MORRA S. (1984) Classification in children: two developmental information-processing models. Int. J. Beh. Devel., 7, 1-20.

NEISSER U. (1976) Knowledge and reality. S.Francisco : Freeman and Co.

NORMAN D.A., SHALLICE T. (1985) Attention to action: willed and automatic control of behavior. In: Davidson R.J., Schwartz G.E., Shapiro D. (eds.), Consciousness and self-regulation: advances in research, vol.IV, N.York : Plenum.

PASCUAL-LEONE J. (1969) Cognitive development and cognitive style: a general psychological integration. Unpublished doctoral thesis, Univ. de Geneve.

PASCUAL-LEONE J. (1970) A mathematical model for the transition rule in Piaget's developmental stages. Acta Psych., 63, 301-345.

PASCUAL-LEONE J. (1978) Compounds, confounds and models in developmental processing: a reply to Trabasso and Foellinger J. Exp. Ch. Psych., 26, 18-40.

PASCUAL-LEONE (1980) Constructive problems for constructive theories: the current relevance of Piaget's work and a critique of information-processing simulation psychology. In: Kluwe R., Spada H. (eds.), Developmental models of thinking. N.York : Academic Press. 263-296.

PASCUAL-LEONE J. (1984) Attention, dialectic and mental effort: toward an organismic theory of life stages. In: Commons M.L., Richards F.A., Armon C. (eds.), Beyond formal operations : late adolescent and adult cognitive development. N.York : Praeger. 182-215.

PASCUAL-LEONE J. (in press). An organismic process model of Witkin's field dependence-independence. In: Globerson T. (ed.), Cognitive style and cognitive development. Norwood, NJ: Ablex.

PASCUAL-LEONE J., GOODMAN D. (1979) Intelligence and experience: a neo-Piagetian approach. Instructional Sc., 8, 301-367.

PASCUAL-LEONE J., GOODMAN D., AMMON P., SUBELMAN I. (1978) Piagetian theory and neo-Piagetian analysis as psychological guides in education. In: Gallagher J., Easley J.A. (eds.), Knowledge and development (vol.2), N.York : Plenum Press. 243-289.

PASCUAL-LEONE J., IJAZ H. (in press). Mental capacity testing as a form of intellectual-developmental assessment. In: Samura R. (ed.), Assessment and placement of minority students. Toronto: Hogrefe International.

PASCUAL-LEONE J., JOHNSON J., GOODMAN D., HAMELUCK D., THEODOR L.H. (1981) Interruption effects in backward pattern masking: the neglected role of fixation stimuli. Proceedings of the 3rd conference of the Cognitive Science Society, 320-323.

PASCUAL-LEONE J., SPARKMAN E. (1980) The dialectics of empiricism and rationalism: a last methodological reply to Trabasso. J. Exp. Ch. Psych., 29, 88-101.

PEW R.W. (1979) Secondary tasks and workload measurement. In: Moray N. (ed.), Mental workload. Its theory and measurement, N.York : Plenum. 23-28.

POSNER M.I., ROSSMAN E. (1965) Effects of size and location of information transforms upon short-term retention. J. Exp. Psych., 70, 496-505.

SALAME' P., BADDELEY A.D. (1982) Disruption of short-term memory by unattended speech: implications for the structure of working memory. J. Verb. Learn. Verb. Beh., 21, 150-164.

SCARDAMALIA M. (1977) Information-processing capacity and the problem of horizontal decalage: a demonstration using combinatorial reasoning tasks. Ch. Dev., 48, 28-37.

SCHOFIELD N.J., ASHMAN A.F. (1986) The relationship between digit span and cognitive processing across ability groups. Intelligence, 10, 59-73.

SHALLICE T. (1982) Specific impairments of planning. Phil. Trans. Roy. Soc. Lond., B-298, 199-209.

STERNBERG R.J. (1981) Testing and cognitive psychology. Amer. Psych., 36, 1181-1189.

VALLAR G., BADDELEY A.D. (1984) Fractionation of working memory: neuropsychological evidence for a phonological short-term store. J. Verb. Learn. Verb. Beh., 23, 151-161.

WITKIN H.A., DYK R.B., FATERSON H.F., GOODENOUGH D.R., KARP S.A. (1962) Psychological differentiation. N.York: Wiley.

Table 1. Factor Analysis

Variables	Factor Loadings		
	F1	F2	F3
Raven's PM47	.78	.20	.19
Goodenough's Draw-A-Man	.57	.34	.24
Figural Intersections	.34	.16	.38
W.I.P.I.	.01	.10	-.34
Forward Digit Span	.11	.57	.28
Backward Digit Span	.38	.29	.40
Corsi Test Forward	.32	.35	.34
Corsi Test Backward	.51	.24	.36
Speed of Articulation	.33	-.49	-.13
Verbal Fluency	.19	.43	.18
Figure Reconstruction	.71	.37	-.05
Counting Span	.28	.42	.48
Block Design	.67	.27	.23
Forward Word Span	.27	.54	.19
Backward Word Span	.30	.39	.35
Mr.Cucumber Test	.38	.08	.53
Vocabulary	.49	.56	.17
eigenvalue	3.5	2.3	1.6
% of total variance	20.6	13.7	9.4

Note: N = 183. The loadings whose absolute value is greater or equal to .35 are underlined.

BEST COPY AVAILABLE

b) Factor 2: Verbal Ability

	FDS	SPA	FWS	VOC	VFL
FDS (Forward Digit Span)		208'	373"	213"	202'
SPA (SPeed of Articulation)	208'		106	348"	202'
FWS (Forward Word Span)	373"	106		251"	221'
VOC (VOCabulary)	213'	348"	251"		300"
VFL (Verbal FLuency)	202'	202'	221'	300"	

c) Factor 3: M Capacity

	BDS	CST	BWS	MCT	FIT
BDS (Backward Digit Span)		289"	280"	231'	251"
CST (Counting Span Test)	289"		332"	249"	210'
BWS (Backward Word Span)	280"	332"		213'	239"
MCT (Mr.Cucumber Test)	231'	249"	213'		313"
FIT (Figural Intersections Test)	251"	210'	239"	313"	

Note: N = 191. Decimal points omitted. The correlations of the Speed of Articulation (a measure of time) have been changed of sign. The significant correlations are marked with ' ($p<.01$), or with " ($p<.001$).

BEST COPY AVAILABLE

Table 3. Correlations of the "spatial" tests of H Capacity with the "verbal" tests of H Capacity and with the tests of Verbal Ability (with age partialled out).

	Verbal tests of H Capacity			tests of Verbal Ability				
	RDS	CSP	FMS	FDS	SPA	FWS	VGC	VFL
MCT	231'	249"	213'	092	001	001	093	143
FIT	251"	210'	239"	164	192'	166	278"	148

Note: N = 191. Decimal points omitted. The correlations of the Speed of Articulation (a measure of time) have been changed of sign. The significant correlations are marked with ' ($p<.01$), or with " ($p<.001$). The meanings of the abbreviations are the same as in tab. 2.

Table 5. Correlations of the "verbal" tests of the Wechsler with the "performance" tests of Wechsler, and with the Field Independence & Spatial Ability test scores (N = 191)

	Verbal tests of Wechsler		Tests of Field Independence & Spatial Ability				
	MCT	CIT	FPC	GLD	BAM	BPM	CIB
SDS	.251'	.251"	.232"	.301"	.217'	.345"	.247"
DSP	.145"	.210'	.144	.279"	.145"	.214'	.140
RMS	.215'	.235"	.229'	.269"	.247"	.236"	.192'

Notes: N = 191. Decimal points omitted. The significant correlations are marked with ' ($p<.01$), or with " ($p<.001$). The meanings of the abbreviations are the same as in tab. 2.

Table 5. Mean scores and analysis of variance for tests of N Experiment 1.

Subjects	Test					
	EDS	ESP	EWS	MCT	FIT	mean
Six	2.90	2.40	3.50	2.20	2.10	2.40
Seven	3.18	2.44	3.68	2.79	3.79	2.86
Eight	3.58	2.92	3.74	2.86	3.64	3.15
Nine	3.86	3.12	3.47	3.25	3.51	3.45
Ten	4.12	3.84	3.84	3.97	4.14	3.98
mean	3.65	3.65	3.39	3.16	3.71	

Factors	Effects
Age	$F(4,186) = 29.06$, $p<.001$
Tests	$F(4,744) = 16.71$, $p<.001$
Age x Tests	$F(16,744) = 1.34$, n.s. $(p>.14)$